Bab 3: Kelas-Kelas Graf Dasar

Pada dua bab pertama, kita terutama bekerja dengan graf tak berarah sederhana. Kita menulis graf sebagai

$$G = (V, E),$$

dengan V himpunan simpul tak kosong dan E himpunan sisi. Untuk graf tak berarah sederhana, setiap sisi adalah pasangan tak berurutan u,v\ dengan u≠ v, dan tidak ada dua sisi yang menghubungkan pasangan simpul yang sama.

Namun data statistik jarang selalu sesederhana itu. Kadang hubungan memiliki arah: seseorang mengirim pesan kepada orang lain, bukan hanya "berhubungan". Kadang dua objek memiliki lebih dari satu jenis hubungan: dua kota dapat terhubung oleh jalan raya, rel kereta, dan rute penerbangan. Kadang hubungan memiliki bobot: jarak, biaya, frekuensi interaksi, korelasi, atau intensitas efek. Kadang simpul terbagi menjadi dua tipe berbeda, seperti pasien dan rumah sakit, penulis dan artikel, atau responden dan pilihan jawaban.

Karena itu, sebelum mempelajari algoritma lebih lanjut, kita harus membedakan beberapa kelas graf dasar. Kelas graf adalah kategori graf berdasarkan aturan tentang sisi, arah, bobot, label, dan tipe simpul. Terminologi dasar seperti graf sederhana, multigraf, gelang, graf berarah, graf berbobot, graf berlabel, dan graf bipartit merupakan bagian standar dari teori graf modern, meskipun beberapa konvensi detail dapat berbeda antarpenulis (Diestel, 2017; West, 2001).

Tujuan bab ini bukan menghafal daftar istilah. Tujuannya adalah memahami bahwa memilih kelas graf berarti memilih asumsi pemodelan. Dalam statistik, asumsi ini menentukan bagaimana data dibaca, bagaimana parameter dimodelkan, dan algoritma apa yang sesuai.

3.1 Mengapa kelas graf penting?

Misalkan kita memiliki data komunikasi antarindividu. Ada beberapa cara untuk mengubah data itu menjadi graf.

Jika kita hanya mencatat apakah dua orang pernah berkomunikasi, maka graf yang wajar adalah graf tak berarah sederhana:

$$\{A, B\} \in E$$

berarti A dan B pernah berkomunikasi setidaknya sekali.

Tetapi jika kita mencatat siapa mengirim pesan kepada siapa, maka arah menjadi penting. Sisi dari A ke B berbeda dari sisi dari B ke A. Ini memerlukan graf berarah.

Jika kita mencatat jumlah pesan, maka hubungan memiliki bobot. Sisi antara A dan B mungkin diberi bobot 37, artinya ada 37 pesan.

Jika kita mencatat setiap pesan sebagai kejadian terpisah, maka bisa ada banyak sisi dari A ke B. Ini mengarah ke multigraf atau graf berarah dengan sisi rangkap.

Jadi, satu sumber data dapat melahirkan beberapa graf berbeda. Perbedaannya bukan hanya teknis. Perbedaannya mengubah pertanyaan statistik.

Sebagai contoh:

- graf sederhana menjawab: “apakah hubungan ada atau tidak?”
- graf berbobot menjawab: “seberapa kuat hubungan itu?”
- graf berarah menjawab: “siapa memengaruhi atau mengirim ke siapa?”
- multigraf menjawab: “berapa banyak kejadian relasional terjadi?”
- graf bipartit menjawab: “objek tipe pertama terhubung ke objek tipe kedua yang mana?”

Dalam analisis jaringan sosial, pembedaan antara relasi tak berarah, berarah, bernilai, dan relasi dua-mode atau bipartit adalah bagian penting dari formulasi data jaringan (Wasserman & Faust, 1994). Dalam ilmu jaringan modern, pilihan antara graf sederhana, berbobot, berarah, dan bipartit juga memengaruhi definisi ukuran jaringan seperti derajat, jarak, centrality, dan clustering (Newman, 2010).

3.2 Graf sederhana

Kita mulai dari kelas yang paling bersih secara matematis.

Sebuah graf sederhana adalah graf tak berarah yang memenuhi dua syarat:

1. tidak memiliki gelang, yaitu sisi dari suatu simpul ke dirinya sendiri;

2. tidak memiliki sisi rangkap, yaitu lebih dari satu sisi yang menghubungkan pasangan simpul yang sama.

Untuk graf sederhana,

$$E \subseteq \{\{u, v\} : u, v \in V,\ u \neq v\}.$$

Artinya, setiap sisi adalah himpunan dua simpul berbeda. Karena u,v\=v,u\, maka sisi tidak memiliki arah.

Contoh 3.1: Graf pertemanan sederhana

Misalkan

$$V = \{A, B, C, D\}$$

adalah empat mahasiswa. Kita definisikan sisi jika dua mahasiswa saling mengenal. Misalnya,

$$E = \{\{A, B\}, \{A, C\}, \{B, D\}\}.$$

Graf ini sederhana karena:

- tidak ada sisi seperti A,A\;
- tidak ada dua salinan sisi A,B\;
- hubungan dianggap simetris.

Derajat setiap simpul adalah:

$$\deg(A) = 2, \quad \deg(B) = 2, \quad \deg(C) = 1, \quad \deg(D) = 1.$$

Jumlah derajatnya adalah

$$2 + 2 + 1 + 1 = 6 = 2|E|.$$

Ini sesuai dengan Handshaking Lemma dari Bab 2.

Contoh 3.2: Graf korelasi antarvariabel

Dalam statistik, graf sederhana sering dipakai untuk merangkum hubungan antarvariabel. Misalkan kita memiliki variabel

$$V = \{X_1, X_2, X_3, X_4\}.$$

Kita dapat membuat sisi antara X_i dan X_j jika korelasi absolut sampelnya melebihi ambang tertentu, misalnya

$$|\widehat{\mathrm{cor}}(X_i, X_j)| > 0.7.$$

Jika hubungan hanya dicatat sebagai “ada” atau “tidak ada”, maka hasilnya adalah graf sederhana.

Namun perlu hati-hati. Graf seperti ini bergantung pada ambang yang dipilih, ukuran sampel, dan ketidakpastian estimasi. Dua korelasi yang tampak berbeda dalam sampel belum tentu berbeda secara statistik. Jadi, graf sederhana dapat menjadi ringkasan awal, tetapi bukan pengganti inferensi statistik.

Kapan graf sederhana cocok?

Graf sederhana cocok ketika:

- relasi bersifat biner: ada atau tidak ada;
- arah relasi tidak penting atau memang simetris;
- intensitas hubungan tidak sedang dianalisis;
- tidak perlu membedakan banyak kejadian antara pasangan simpul yang sama.

Contoh:

- peta ketetanggaan wilayah: dua wilayah bertetangga atau tidak;
- konflik jadwal: dua mata kuliah berbenturan atau tidak;
- jaringan kolaborasi: dua penulis pernah menulis artikel bersama atau tidak;
- struktur ketergantungan awal: dua variabel dianggap terkait atau tidak.

Graf sederhana sering menjadi titik awal karena definisinya bersih dan banyak teorema dasar graf dirumuskan untuk kelas ini (Diestel, 2017; West, 2001).

3.3 Multigraf: ketika sisi rangkap bermakna

Dalam graf sederhana, dua simpul hanya boleh dihubungkan oleh paling banyak satu sisi. Tetapi dalam banyak data, dua objek dapat memiliki beberapa hubungan berbeda atau beberapa kejadian hubungan.

Sebuah multigraf adalah graf yang mengizinkan sisi rangkap. Sisi rangkap adalah dua atau lebih sisi yang memiliki ujung simpul yang sama.

Secara intuitif, jika A dan B dihubungkan oleh tiga sisi, maka ada tiga hubungan atau tiga kejadian antara A dan B.

Ada beberapa cara formal untuk mendefinisikan multigraf. Salah satu cara yang rapi adalah memisahkan himpunan sisi dari pasangan ujungnya. Kita dapat menulis

$$G = (V, E, \phi),$$

dengan φ sebagai fungsi insidensi yang memetakan setiap sisi ke pasangan simpul ujungnya. Misalnya,

$$\phi(e_1) = \{A, B\}, \quad \phi(e_2) = \{A, B\}, \quad \phi(e_3) = \{A, B\}.$$

Di sini e_1, e_2, e_3 adalah tiga sisi berbeda, walaupun semuanya menghubungkan A dan B. Formulasi semacam ini diperlukan karena jika E hanya dianggap himpunan pasangan simpul, maka A,B\ tidak dapat muncul tiga kali. Buku teori graf sering membedakan graf sederhana dan multigraf dengan perhatian terhadap sisi rangkap dan gelang, walaupun notasi formalnya dapat bervariasi (West, 2001).

Contoh 3.3: Transaksi berulang

Misalkan simpul adalah akun bank:

$$V = \{A, B, C\}.$$

Selama satu hari, terjadi transaksi:

- A mengirim ke B pada pukul 09.00;
- A mengirim ke B pada pukul 11.00;
- B mengirim ke C pada pukul 13.00.

Jika arah diabaikan, kita dapat membuat multigraf tak berarah dengan sisi

$$e_1 : \{A, B\}, \quad e_2 : \{A, B\}, \quad e_3 : \{B, C\}.$$

Sisi e_1 dan e_2 adalah sisi rangkap.

Jika kita mereduksi data ini menjadi graf sederhana, maka dua transaksi antara A dan B hanya menjadi satu sisi. Informasi frekuensi hilang. Kadang reduksi itu berguna, tetapi kadang justru merusak pertanyaan penelitian.

Derajat dalam multigraf

Dalam multigraf tak berarah tanpa gelang, derajat simpul tetap dihitung sebagai banyak sisi yang bersisian dengannya, dengan sisi rangkap dihitung satu per satu.

Pada contoh di atas:

$$\deg(A) = 2, \quad \deg(B) = 3, \quad \deg(C) = 1.$$

Jumlah derajatnya adalah

$$2 + 3 + 1 = 6 = 2|E|.$$

Karena $|E|=3$, Handshaking Lemma tetap berlaku:

$$\sum_{v \in V} \deg(v) = 2|E|.$$

Alasannya sama seperti sebelumnya: setiap sisi memiliki dua ujung, dan setiap ujung menyumbang satu ke derajat simpul terkait.

Kapan multigraf cocok?

Multigraf cocok ketika kejadian ganda tidak boleh dilebur menjadi satu relasi biner. Contohnya:

- beberapa transaksi antara dua akun;
- beberapa perjalanan antara dua kota;
- beberapa kolaborasi antara dua peneliti;

- beberapa interaksi sosial antara dua orang dalam periode pengamatan;
- beberapa jenis hubungan antara pasangan unit yang sama.

Dalam statistik, multigraf sering muncul ketika data asli berbentuk daftar kejadian. Jika setiap baris data adalah satu kejadian relasional, maka menggabungkannya menjadi graf sederhana adalah keputusan pemodelan, bukan langkah otomatis.

3.4 Graf dengan gelang

Sebuah gelang atau loop adalah sisi yang menghubungkan sebuah simpul dengan dirinya sendiri.

Jika v∈ V, maka gelang pada v dapat ditulis sebagai

$$\{v, v\}$$

untuk graf tak berarah, atau sebagai

$$(v, v)$$

untuk graf berarah.

Graf sederhana tidak mengizinkan gelang. Tetapi beberapa model membutuhkan gelang.

Contoh 3.4: Transisi diri

Misalkan simpul menyatakan keadaan cuaca:

$$V = \{\text{cerah}, \text{hujan}, \text{berawan}\}.$$

Jika kita mencatat transisi cuaca dari hari ke hari, maka mungkin terjadi:

$$\text{cerah} \to \text{cerah}.$$

Ini adalah transisi dari suatu keadaan ke keadaan yang sama. Dalam graf berarah, ini dapat direpresentasikan sebagai gelang pada simpul “cerah”.

Contoh seperti ini sering muncul dalam model transisi, termasuk rantai Markov, meskipun teori rantai Markov memiliki notasi probabilistik khusus sendiri. Graf di sini membantu menggambarkan transisi yang mungkin.

Contoh 3.5: Relasi diri dalam data

Misalkan simpul adalah halaman web. Jika sebuah halaman memiliki tautan ke dirinya sendiri, maka ada sisi berarah dari halaman itu ke dirinya sendiri. Itu adalah gelang.

Namun dalam banyak konteks statistik, gelang tidak bermakna. Misalnya, jika simpul adalah individu dan sisi berarti “dua individu berbeda adalah teman”, maka sisi A,A\ tidak masuk akal. Karena itu graf pertemanan biasanya dianggap tanpa gelang.

Derajat dan gelang

Dalam graf tak berarah, konvensi standar adalah bahwa satu gelang menyumbang 2 terhadap derajat simpulnya. Alasannya sesuai dengan gagasan “ujung sisi”: sebuah gelang memiliki dua ujung, tetapi kedua ujung itu berada pada simpul yang sama. Konvensi ini menjaga kebenaran Handshaking Lemma untuk graf tak berarah yang memiliki gelang (West, 2001).

Misalkan

$$V = \{A, B\}$$

dan

$$E = \{\{A, A\}, \{A, B\}\}.$$

Ada satu gelang pada A dan satu sisi antara A dan B. Maka

$$\deg(A) = 3, \quad \deg(B) = 1.$$

Mengapa deg(A)=3? Karena gelang menyumbang 2, dan sisi A,B\ menyumbang 1.

Jumlah derajat:

$$3 + 1 = 4 = 2|E|.$$

Karena $|E|=2$, Handshaking Lemma tetap berlaku.

Pseudograf

Beberapa buku menggunakan istilah pseudograf untuk graf yang mengizinkan gelang dan sisi rangkap. Namun istilah ini tidak selalu dipakai secara seragam. Karena itu, dalam buku ini kita akan lebih sering mengatakan secara eksplisit: “graf dengan gelang”, “multigraf”, atau “multigraf dengan gelang”.

3.5 Graf berarah atau digraf

Sejauh ini, sisi biasanya dianggap tidak memiliki arah. Tetapi banyak relasi tidak simetris.

Jika A mengikuti B di media sosial, belum tentu B mengikuti A. Jika artikel P mengutip artikel Q, maka arah kutipan adalah dari P ke Q, bukan sebaliknya. Jika variabel X diasumsikan menyebabkan variabel Y, arah $X \to Y$ tidak dapat ditukar begitu saja.

Sebuah graf berarah, atau digraf, adalah graf yang sisinya memiliki arah. Sisi berarah sering disebut busur atau arc.

Secara formal, graf berarah dapat ditulis sebagai

$$D = (V, A),$$

dengan V himpunan simpul dan A himpunan busur. Setiap busur adalah pasangan berurutan

$$(u, v),$$

dengan $u,v \in V$. Busur (u,v) berarti ada sisi dari u menuju v. Simpul u disebut ekor atau asal busur, sedangkan v disebut kepala atau tujuan busur.

Karena pasangan berurutan memperhatikan urutan, maka

$$(u,v) \neq (v,u)$$

kecuali jika u=v.

Contoh 3.6: Jaringan pengiriman pesan

Misalkan

$$V = \{A,B,C\}$$

dan busur

$$A = \{(A,B),(B,A),(A,C)\}.$$

Di sini:

- A mengirim ke B;
- B mengirim ke A;
- A mengirim ke C.

Busur (A,B) dan (B,A) adalah dua busur berbeda. Keduanya bersama-sama menunjukkan komunikasi dua arah, tetapi secara matematis tetap dua relasi berarah.

Derajat masuk dan derajat keluar

Dalam graf berarah, derajat biasa dipecah menjadi dua.

Derajat keluar dari simpul v, ditulis

$$\deg^+(v),$$

adalah banyak busur yang keluar dari v.

Derajat masuk dari simpul v, ditulis

$$\deg^-(v),$$

adalah banyak busur yang masuk ke v.

Pada contoh

$$A = \{(A, B), (B, A), (A, C)\},$$

kita punya:

$$\deg^+(A) = 2, \quad \deg^-(A) = 1,$$

karena A memiliki dua busur keluar, yaitu (A,B) dan (A,C), serta satu busur masuk, yaitu (B,A).

Untuk simpul lain:

$$\deg^+(B) = 1, \quad \deg^-(B) = 1,$$

dan

$$\deg^+(C) = 0, \quad \deg^-(C) = 1.$$

Dalam graf berarah hingga tanpa menghitung gelang secara khusus, setiap busur menyumbang satu ke total derajat keluar dan satu ke total derajat masuk. Karena itu,

$$\sum_{v \in V} \deg^+(v) = |A|$$

dan

$$\sum_{v \in V} \deg^-(v) = |A|.$$

Akibatnya,

$$\sum_{v \in V} \deg^+(v) = \sum_{v \in V} \deg^-(v).$$

Ini adalah analog berarah dari Handshaking Lemma. Identitas dasar derajat masuk dan keluar seperti ini merupakan bagian standar dari teori digraf (Diestel, 2017; West, 2001).

Sumber dan muara

Dalam graf berarah, dua istilah sederhana sering berguna.

Sebuah simpul disebut sumber jika derajat masuknya nol:

$$\deg^-(v) = 0.$$

Artinya, tidak ada busur yang masuk ke simpul tersebut.

Sebuah simpul disebut muara jika derajat keluarnya nol:

$$\deg^+(v) = 0.$$

Artinya, tidak ada busur yang keluar dari simpul tersebut.

Pada contoh sebelumnya, simpul C adalah muara karena

$$\deg^+(C) = 0.$$

Simpul C menerima busur dari A, tetapi tidak mengirim busur ke mana pun.

Contoh 3.7: Graf kausal sederhana

Misalkan kita memiliki tiga variabel:

$$V = \{\text{pendidikan}, \text{pendapatan}, \text{kesehatan}\}.$$

Sebuah model substantif mungkin mengasumsikan:

$$\text{pendidikan} \to \text{pendapatan}$$

dan

$$\text{pendapatan} \to \text{kesehatan}.$$

Graf ini bukan sekadar hubungan simetris. Arah membawa interpretasi: pendidikan diposisikan sebagai pendahulu pendapatan, dan pendapatan sebagai pendahulu kesehatan.

Dalam analisis kausal modern, graf berarah asiklik atau DAG digunakan untuk menyatakan asumsi struktural tentang hubungan sebab-akibat dan independensi kondisional. Kita akan membahas ini secara serius pada Bab 21. Untuk sekarang, yang penting adalah memahami bahwa arah pada sisi adalah bagian dari model, bukan hiasan visual.

Kapan graf berarah cocok?

Graf berarah cocok ketika relasi memiliki orientasi alami, misalnya:

- pengirim dan penerima pesan;
- pengikut dan diikuti;
- halaman web dan tautan keluar;
- artikel yang mengutip dan artikel yang dikutip;
- alur kerja;
- dependensi perangkat lunak;
- asumsi kausal antarvariabel;
- transisi antarstatus.

Dalam data jaringan sosial dan jaringan informasi, arah sering mengubah kesimpulan. Seseorang dengan banyak busur masuk mungkin populer atau sering dikutip, sedangkan seseorang dengan banyak busur keluar mungkin aktif mencari atau menyebarkan koneksi. Newman menekankan bahwa jaringan berarah dan tak berarah sering membutuhkan ukuran dan interpretasi yang berbeda (Newman, 2010).

3.6 Graf berbobot dan graf tak berbobot

Pada graf sederhana, sisi hanya menyatakan ada atau tidak ada hubungan. Tetapi sering kali kita ingin mencatat kekuatan, jarak, biaya, probabilitas, frekuensi, risiko, atau ukuran lain pada sisi.

Sebuah graf berbobot adalah graf yang setiap sisinya diberi nilai numerik, disebut bobot. Secara formal, selain graf G=(V,E), kita memiliki fungsi bobot

$$w : E \to \mathbb{R}.$$

Untuk setiap sisi $e \in E$, nilai $w(e)$ adalah bobot sisi tersebut.

Jika graf tidak memiliki bobot, kita menyebutnya graf tak berbobot. Graf tak berbobot dapat dipandang sebagai graf yang hanya mencatat struktur relasi, bukan intensitas relasi.

Contoh 3.8: Jarak antarwilayah

Misalkan simpul adalah kota:

$$V = \{\text{Bandung}, \text{Jakarta}, \text{Cirebon}\}.$$

Sisi menyatakan ada rute langsung, dan bobot menyatakan jarak kilometer:

$$w(\{\text{Bandung}, \text{Jakarta}\}) = 150,$$

$$w(\{\text{Bandung}, \text{Cirebon}\}) = 130,$$

$$w(\{\text{Jakarta}, \text{Cirebon}\}) = 220.$$

Di sini bobot lebih besar berarti jarak lebih jauh. Dalam masalah jarak terpendek, bobot seperti ini sering ditafsirkan sebagai biaya yang ingin diminimalkan.

Contoh 3.9: Frekuensi interaksi

Misalkan simpul adalah responden survei. Sisi antara dua responden diberi bobot jumlah percakapan dalam satu minggu.

Jika

$$w(\{A, B\}) = 12$$

dan

$$w(\{A, C\}) = 2,$$

maka hubungan A-B lebih intens daripada hubungan A-C, setidaknya menurut ukuran frekuensi percakapan.

Contoh 3.10: Korelasi sebagai bobot

Misalkan simpul adalah variabel statistik. Kita dapat memberi bobot pada sisi antara X_i dan X_j menggunakan korelasi sampel:

$$w(\{X_i, X_j\}) = \widehat{\mathrm{cor}}(X_i, X_j).$$

Dalam kasus ini bobot dapat bernilai negatif. Ini berbeda dari jarak atau frekuensi, yang biasanya tidak negatif.

Karena itu, kata “bobot” tidak selalu berarti hal yang sama. Bobot dapat berupa:

- jarak;
- biaya;
- kapasitas;
- frekuensi;
- kekuatan hubungan;
- korelasi;
- kemiripan;
- risiko;
- probabilitas;
- estimasi parameter.

Interpretasi bobot harus dinyatakan dengan jelas. Dua graf dapat sama-sama berbobot, tetapi algoritma yang sesuai bisa berbeda. Misalnya, algoritma jarak terpendek biasanya menganggap bobot sebagai biaya, sedangkan analisis asosiasi dapat menganggap bobot sebagai kekuatan hubungan. Pembahasan jaringan berbobot sebagai perluasan dari jaringan tak berbobot merupakan tema umum dalam ilmu jaringan (Newman, 2010).

Bobot bukan otomatis probabilitas

Dalam statistik, kita harus sangat hati-hati: bobot tidak otomatis berarti probabilitas.

Jika

$$w(e) = 0.8,$$

maka nilai ini bisa berarti korelasi 0.8, probabilitas 0.8, kemiripan 0.8, atau skor normalisasi 0.8. Tanpa definisi, angka itu tidak memiliki interpretasi statistik yang jelas.

Sebagai aturan kerja:

> Selalu nyatakan apa arti bobot, satuannya, rentangnya, dan apakah bobot yang lebih besar berarti hubungan lebih kuat, lebih dekat, lebih mahal, atau lebih berisiko.

3.7 Graf berlabel dan graf tak berlabel

Dalam graf yang kita gunakan untuk data, simpul biasanya memiliki nama. Misalnya:

$$V = \{\mathrm{Ani}, \mathrm{Budi}, \mathrm{Citra}\}.$$

Nama-nama ini adalah label simpul. Sebuah graf berlabel adalah graf yang simpulnya, sisinya, atau keduanya memiliki label yang membedakan identitasnya.

Label simpul dapat berupa:

- nama responden;
- kode wilayah;
- ID pasien;
- nama variabel;
- koordinat lokasi;
- jenis kelompok;
- atribut demografis.

Label sisi dapat berupa:

- jenis hubungan;
- waktu kejadian;
- kategori transaksi;
- sumber data;
- nilai pengukuran;
- tanda positif atau negatif.

Sebaliknya, dalam graf tak berlabel, yang diperhatikan hanya pola keterhubungan abstrak, bukan nama simpulnya.

Contoh 3.11: Dua graf dengan pola sama

Pertimbangkan dua graf:

$$G_1: \quad V_1 = \{A, B, C\}, \quad E_1 = \{\{A, B\}, \{B, C\}\}.$$

$$G_2: \quad V_2 = \{X, Y, Z\}, \quad E_2 = \{\{X, Y\}, \{Y, Z\}\}.$$

Jika label diperhatikan, maka G_1 dan G_2 berbeda karena simpulnya bernama berbeda. Tetapi secara bentuk, keduanya sama: masing-masing adalah lintasan dengan tiga simpul.

Dalam teori graf, dua graf yang memiliki bentuk sama meskipun label berbeda disebut isomorfik. Secara informal, isomorfisme adalah pencocokan ulang nama simpul yang mempertahankan sisi. Konsep ini akan muncul lagi ketika kita membahas motif, struktur jaringan, dan masalah komputasi.

Label dalam statistik

Dalam statistik terapan, label sering sangat penting. Jika simpul adalah pasien, maka menukar label pasien dapat mengubah arti data. Jika simpul adalah wilayah, label wilayah membawa informasi geografis, sosial, dan administratif.

Namun dalam beberapa analisis, kita sengaja mengabaikan label. Misalnya, saat membandingkan bentuk jaringan kecil atau motif, kita mungkin hanya peduli bahwa suatu pola berbentuk segitiga, bintang, atau lintasan, bukan siapa tepatnya simpulnya. Dalam analisis jaringan, pembedaan antara struktur relasional dan atribut aktor merupakan bagian penting dari cara data jaringan didefinisikan (Wasserman & Faust, 1994).

Contoh 3.12: Label sisi sebagai jenis relasi

Misalkan simpul adalah organisasi:

$$V = \{O_1, O_2, O_3\}.$$

Antara O_1 dan O_2, mungkin ada dua jenis relasi:

- kerja sama riset;

- pendanaan.

Jika kita ingin menyimpan jenis relasi, kita dapat memberi label pada sisi, misalnya:

$$\ell(e_1) = \text{kerja sama riset},$$

$$\ell(e_2) = \text{pendanaan}.$$

Jika kedua relasi terjadi antara pasangan organisasi yang sama, maka kita dapat memodelkannya sebagai multigraf berlabel, atau sebagai graf dengan sisi yang memiliki atribut kategori. Pilihan ini bergantung pada pertanyaan analisis dan format komputasi yang digunakan.

3.8 Graf bipartit

Sekarang kita masuk ke kelas graf yang sangat penting dalam statistik data relasional.

Sebuah graf disebut bipartit jika himpunan simpulnya dapat dibagi menjadi dua bagian tak beririsan,

$$V = U \cup W, \quad U \cap W = \varnothing,$$

sehingga setiap sisi menghubungkan satu simpul di U dengan satu simpul di W. Tidak ada sisi antara dua simpul di U, dan tidak ada sisi antara dua simpul di W.

Dengan kata lain, semua sisi "menyeberang" dari satu bagian ke bagian lain.

Contoh 3.13: Responden dan produk

Misalkan

$$U = \{\text{Ani}, \text{Budi}, \text{Citra}\}$$

adalah responden, dan

$$W = \{\text{Produk } 1, \text{Produk } 2, \text{Produk } 3\}$$

adalah produk.

Sisi menyatakan bahwa seorang responden membeli suatu produk. Misalnya,

$$E = \{\{\text{Ani}, \text{Produk } 1\}, \{\text{Ani}, \text{Produk } 3\}, \{\text{Budi}, \text{Produk } 2\}, \{\text{Citra}, \text{Produk } 1\}\}.$$

Ini adalah graf bipartit karena sisi selalu menghubungkan responden dengan produk. Tidak ada sisi langsung antara Ani dan Budi, dan tidak ada sisi langsung antara Produk 1 dan Produk 2.

Graf seperti ini sering disebut jaringan dua-mode atau jaringan afiliasi dalam literatur jaringan sosial, karena simpul berasal dari dua tipe entitas yang berbeda (Wasserman & Faust, 1994).

Contoh 3.14: Penulis dan artikel

Misalkan U adalah himpunan penulis dan W adalah himpunan artikel. Kita hubungkan penulis u∈ U ke artikel w∈ W jika penulis tersebut menulis artikel itu.

Jika sebuah artikel memiliki tiga penulis, maka artikel tersebut terhubung ke tiga simpul penulis.

Graf bipartit ini menyimpan struktur kolaborasi secara lebih lengkap daripada graf proyeksi penulis-penulis. Jika kita memproyeksikan menjadi graf penulis, dua penulis dihubungkan jika pernah menulis artikel bersama. Tetapi proyeksi itu dapat kehilangan informasi: apakah tiga penulis menulis satu artikel bersama, atau setiap pasangan menulis artikel berbeda?

Graf bipartit lengkap

Graf bipartit lengkap dengan bagian berukuran m dan n ditulis

$$K_{m,n}.$$

Graf ini memiliki dua bagian:

$$|U| = m, \quad |W| = n,$$

dan setiap simpul di U terhubung ke setiap simpul di W.

Jumlah sisinya adalah

$$mn.$$

Mengapa? Karena untuk setiap dari m simpul di U, ada n pilihan simpul di W. Jadi total pasangan lintas bagian adalah m n.

Sebagai contoh, K(2,3) memiliki

$$2 \cdot 3 = 6$$

sisi.

Contoh 3.15: Desain eksperimen sebagai graf bipartit

Graf bipartit dapat membantu menggambarkan desain eksperimen atau desain observasional.

Misalkan U adalah unit eksperimen dan W adalah perlakuan. Sisi u,w\ berarti unit u menerima perlakuan w.

Dalam desain sederhana, setiap unit mungkin terhubung ke tepat satu perlakuan. Dalam desain yang lebih kompleks, unit dapat menerima beberapa perlakuan atau kombinasi perlakuan. Representasi bipartit membantu melihat apakah semua perlakuan cukup terwakili, apakah ada unit yang tidak mendapat perlakuan, atau apakah struktur penugasan tidak seimbang.

Proyeksi bipartit

Dari graf bipartit, kita sering membuat graf satu-mode melalui proyeksi.

Jika graf bipartit menghubungkan responden dengan produk, maka ada dua proyeksi alami:

1. proyeksi responden: dua responden dihubungkan jika membeli produk yang sama;
2. proyeksi produk: dua produk dihubungkan jika dibeli oleh responden yang sama.

Proyeksi ini berguna, tetapi harus hati-hati. Proyeksi dapat membuat graf menjadi jauh lebih padat dan dapat menyembunyikan struktur asli dua-mode. Dalam analisis jaringan, jaringan bipartit dan proyeksinya sering memiliki interpretasi yang berbeda, sehingga tidak boleh dianggap saling menggantikan tanpa pertimbangan substantif (Newman, 2010; Wasserman & Faust, 1994).

3.9 Kelas-kelas graf dapat digabungkan

Kelas graf bukan kotak yang saling terpisah. Sebuah graf dapat sekaligus berarah, berbobot, berlabel, dan bipartit.

Misalnya, graf transaksi antara pelanggan dan toko dapat berupa:

- bipartit: simpul pelanggan dan simpul toko;
- berarah: pelanggan membayar ke toko;
- berbobot: bobot sisi adalah nilai transaksi;
- multigraf: satu pelanggan dapat bertransaksi berkali-kali dengan toko yang sama;
- berlabel: setiap sisi memiliki waktu transaksi dan kategori barang.

Secara matematis, struktur seperti ini lebih kaya daripada graf sederhana. Secara statistik, ia juga membawa lebih banyak informasi.

Namun semakin kaya struktur graf, semakin banyak keputusan yang harus dibuat:

- Apakah kejadian berulang disimpan sebagai sisi rangkap atau dijumlahkan menjadi bobot?
- Apakah arah penting?
- Apakah bobot bersifat kontinu, diskret, positif, negatif, atau kategori?
- Apakah simpul memiliki tipe berbeda?
- Apakah gelang bermakna?
- Apakah label harus dipertahankan atau dianonimkan?

Tidak ada jawaban universal. Jawaban bergantung pada pertanyaan ilmiah, desain data, dan metode analisis.

Contoh 3.16: Tiga representasi dari data yang sama

Misalkan data mentah mencatat pesan:

waktu	pengirim	penerima
08.00	A	B
08.05	A	B
08.10	B	A
08.20	A	C

Dari data ini, kita dapat membuat beberapa graf.

Representasi 1: graf sederhana tak berarah

$$E = \{\{A, B\}, \{A, C\}\}.$$

Interpretasi: pasangan mana pernah berkomunikasi?

Representasi 2: graf berarah sederhana

$$A = \{(A, B), (B, A), (A, C)\}.$$

Interpretasi: arah komunikasi mana pernah terjadi?

Representasi 3: graf berarah berbobot

$$w(A, B) = 2, \quad w(B, A) = 1, \quad w(A, C) = 1.$$

Interpretasi: berapa banyak pesan dikirim dari satu orang ke orang lain?

Ketiganya sah, tetapi menjawab pertanyaan berbeda. Jika kita tertarik pada keberadaan koneksi sosial, representasi pertama mungkin cukup. Jika kita tertarik pada ketimpangan komunikasi, representasi ketiga lebih informatif.

3.10 Implikasi statistik dari pilihan kelas graf

Dalam statistik, graf bukan hanya gambar. Graf adalah struktur data dan struktur model. Karena itu, kelas graf yang dipilih memengaruhi inferensi.

Graf sederhana dan data biner

Jika graf sederhana digunakan, maka setiap kemungkinan sisi biasanya dipandang sebagai variabel biner:

$$Y_{ij} = \begin{cases} 1, & \text{jika sisi antara } i \text{ dan } j \text{ ada,} \\ 0, & \text{jika tidak ada.} \end{cases}$$

Ini cocok untuk pertanyaan seperti:

- apakah dua orang pernah berinteraksi?
- apakah dua wilayah berbatasan?
- apakah dua variabel memiliki asosiasi kuat menurut aturan tertentu?

Model statistik untuk graf sederhana sering berurusan dengan peluang kemunculan sisi, pola ketergantungan antar sisi, atau struktur komunitas. Ini akan muncul lagi ketika kita membahas graf acak dan model statistik jaringan.

Multigraf dan data hitungan

Jika sisi rangkap disimpan, maka data relasional sering berbentuk hitungan kejadian. Misalnya, jumlah pesan dari i ke j, jumlah perjalanan antara dua kota, atau jumlah transaksi antara dua akun.

Dalam kasus seperti ini, mereduksi data menjadi biner dapat membuang variasi penting. Misalnya, pasangan yang bertransaksi sekali dan pasangan yang bertransaksi seribu

Document information

Bab 3: Kelas-Kelas Graf Dasar

Project	Teori Graf untuk Statistik
Document	Document 1.7
Author	Harizahayu
Verifier	Not verified
Downloaded	July 08, 2026 10:36 KST
Status	Working
Document link	https://theorytrace.com/projects/teori-graf-untuk-statistik/documents/bab-3-kelas-kelas--graf-dasar/